(11) Japanese Patent Laid-Open No. 22299/1997

(43) Laid-Open Date: January 21, 1997

(21) Application No. 194092/1995

(22) Application Date: July 7, 1995

(71) Applicant: Kokusai Electric Co. Ltd.

(72) Inventors: Ken Satoh

Request for examination: Not made

(54) [Title of the Invention] VOICE ENCODING COMMUNICATION SCHEME

(57) [Abstract]

[Problem] In the communication by an analysis compositing type voice encoding scheme, reception reproduced voice quality is relieved from deteriorating when narrowing the band at low rate.

[Means for Resolution] On the transmission end, an LSP analyzer 1 carries out LSP-analysis on an input voice signal based on a frame to extract a spectrum parameter (SP). An analyzing filter 2 calculates a sound-source signal removed of SP from the input. A pitch extractor 3 extracts voiced/unvoiced information and pitch period of the sound-source signal, and an encoder 10 encode and sends them to a multiplexer 13. Voiced sound-source and unvoiced sound-source codebooks 8, 9 and an LSP codebook 12 are provided. The

sound-source signal is transformed into a frequency range by a DCT 4 and switched for a voiced and an unvoiced. From the respective codebook 8, 9, corresponding codes are red out and quantized, and sent to the multiplexer 13. A code corresponding to the SP is read out of the codebook 12 and quantized, and sent to the multiplexer 13. The multiplexer 13 multiplexes and delivers them onto the transmission line.

[Claim]

[Claim 1] A voice encoding communication scheme comprising:

on a transmission end, a spectrum parameter analyzer for extracting a spectrum parameter from a voice signal; an analyzing filter for calculating a sound-source signal removed of the spectrum parameter from the voice signal; a pitch extractor for detecting voiced/unvoiced information and a pitch period from the sound-source signal; an encoder for encoding the voiced/unvoiced information and the pitch period; a discrete cosine transformer for transforming the sound-source signal from a time region into a frequency region; a switch for switch-output an output signal of the discrete cosine transformer to any of a voiced output and an unvoiced output according to the voiced/unvoiced information determined by the pitch extractor; a voiced sound-source codebook holding a sound-source signal transformed from a time

region into a frequency region of a voiced sound-source signal previously determined by a training signal; an unvoiced sound-source codebook holding a sound-source signal transformed from a time region into a frequency region of an unvoiced sound-source signal previously determined by a training signal; a voiced sound-source signal quantizer for collating between a voiced output of the switch and a content of the voiced sound-source codebook to replace into a sound-source signal in a corresponding frequency region; an unvoiced sound-source signal quantizer for collating between an unvoiced output of the switch and a content of the unvoiced sound-source codebook to replace into a sound-source signal in a corresponding frequency region; a spectrum parameter codebook holding a parameter previously determined in the spectrum parameter analyzer by a training signal; a spectrum parameter quantizer for collating between a spectrum parameter from the spectrum parameter analyzer and a content of the spectrum parameter codebook to replace into a corresponding spectrum parameter; and a multiplexer for multiplexing together a voiced or unvoiced sound-source signal quantized in the voiced sound-source signal quantizer or in the unvoiced sound-source signal quantizer, a spectrum parameter quantized in the spectrum parameter quantizer, voiced/unvoiced information encoded in the encoder and a pitch period to deliver the same onto a transmission line; and

on a reception end, a separator for receiving a multiplex signal delivered from the multiplexer through the transmission line, and separating and outputting with a quantized voiced or unvoiced sound-source signal, a quantized spectrum parameter, encoded voiced/unvoiced information and a pitch period; a decoder for decoding the separated voiced/unvoiced information and pitch period; a switch for switch-outputting a separated quantixed voiced or unvoiced sound-source signal to any of a voiced sound-source output or an unvoiced sound-source output according to voiced/unvoiced information outputted from the decoder; a spectrum parameter codebook searcher for collating the separated spectrum parameter with a content of a spectrum parameter codebook having a same content as the transmission-end spectrum parameter codebook to replace into a corresponding spectrum parameter; a voiced sound-source codebook searcher for collating a voiced sound-source output of the switch with a content of a voiced sound-source codebook having a same content as the transmission-end voiced sound-source codebook to replace into a voiced sound-source signal in a corresponding frequency region; an unvoiced sound-source codebook searcher for collating an unvoiced sound-source output of the switch with a content of an unvoiced sound-source codebook having a same content as the transmission-end unvoiced sound-source codebook to replace into an unvoiced sound-source signal in a corresponding

frequency region; an inverted discrete cosine transformer for restoring to a time region a voiced or unvoiced sound-source signal in a frequency region outputted from the voiced sound-source codebook searcher and the unvoiced sound-source codebook searcher; a voice reproducer for adding an output of the inverted discrete cosine transformer with pitch period information from the decoder to reproduce a sound source when the output is voiced; and a combining filter for outputting a reproduced voice signal due to outputs of the sound-source reproducer and the spectrum parameter codebook searcher.

[Detailed Description of the Invention]

[0001]

[Technical Field of the Invention]

The present invention relates to a voice encoding communication scheme and, more particularly, to a voice encoding communication scheme by an analysis combining encoding communication method.

[0002]

[Prior Art]

In the voice encoding scheme at a transmission rate of approximately several kbps, favorable reproduced voice quality is provided by the hybrid encoding scheme as represented by a CELP (code excited linear prediction) encoding scheme. However, in an extremely low rate region of 2.4 kbps or lower, the analysis combining encoding scheme is on the

mainstream because of limitation in transmittable information amount.

[0003]

[Problem that the Invention is to Solve]

However, generally, the analysis combining encoding scheme is poor in voice quality because of its scheme to replace a sound-source with a simple model. It is an object of the present invention to provide a voice encoding communication scheme which is improved in sound-source modeling method to solve the deterioration of quality in order to relieve the deterioration in voice quality to be caused in an extremely low rate region in the analysis combining encoding scheme, as a defect in the conventional scheme, i.e. deterioration in voice quality due to simple modeling of the sound source.

[0004]

[Means for Solving the Problem]

A voice encoding communication scheme of the present invention comprises: on a transmission end, a spectrum parameter analyzer for extracting a spectrum parameter from a voice signal; an analyzing filter for calculating a sound-source signal removed of the spectrum parameter from the voice signal; a pitch extractor for detecting voiced/unvoiced information and a pitch period from the sound-source signal; an encoder for encoding the voiced/unvoiced information and the pitch period; a discrete cosine transformer for

transforming the sound-source signal from a time region into a frequency region; a switch for switch-output an output signal of the discrete cosine transformer to any of a voiced output and an unvoiced output according to the voiced/unvoiced information determined by the pitch extractor; a voiced sound-source codebook holding a sound-source signal transformed from a time region into a frequency region of a voiced sound-source signal previously determined by a training signal; an unvoiced sound-source codebook holding a sound-source signal transformed from a time region into a frequency region of an unvoiced sound-source signal previously determined by a training signal; a voiced sound-source signal quantizer for collating between a voiced output of the switch and a content of the voiced sound-source codebook to replace into a sound-source signal in a corresponding frequency region; an unvoiced sound-source signal quantizer for collating between an unvoiced output of the switch and a content of the unvoiced sound-source codebook to replace into a sound-source signal in a corresponding frequency region; a spectrum parameter codebook holding a parameter previously determined in the spectrum parameter analyzer by a training signal; a spectrum parameter quantizer for collating between a spectrum parameter from the spectrum parameter analyzer and a content of the spectrum parameter codebook to replace into a corresponding spectrum parameter; and a multiplexer for

multiplexing together a voiced or unvoiced sound-source signal quantized in the voiced sound-source signal quantizer or in the unvoiced sound-source signal quantizer, a spectrum parameter quantized in the spectrum parameter quantizer, voiced/unvoiced information encoded in the encoder and a pitch period to deliver the same onto a transmission line; and on a reception end, a separator for receiving a multiplex signal delivered from the multiplexer through a transmission line, and separating and outputting with a quantized voiced or unvoiced sound-source signal, a quantized spectrum parameter, encoded voiced/unvoiced information and a pitch period; a decoder for decoding the separated voiced/unvoiced information and pitch period; a switch for switch-outputting a separated quantixed voiced or unvoiced sound-source signal to any of a voiced sound-source output or an unvoiced sound-source output according to voiced/unvoiced information outputted from the decoder; a spectrum parameter codebook searcher for collating the separated spectrum parameter with a content of a spectrum parameter codebook having a same content as the transmission-end spectrum parameter codebook to replace into a corresponding spectrum parameter; a voiced sound-source codebook searcher for collating a voiced sound-source output of the switch with a content of a voiced sound-source codebook having a same content as the transmission-end voiced sound-source codebook to replace into a voiced sound-source

signal in a corresponding frequency region; an unvoiced sound-source codebook searcher for collating an unvoiced sound-source output of the switch with a content of an unvoiced sound-source codebook having a same content as the transmission-end unvoiced sound-source codebook to replace into an unvoiced sound-source signal in a corresponding frequency region; an inverted discrete cosine transformer for restoring to a time region a voiced or unvoiced sound-source signal in a frequency region outputted from the voiced sound-source codebook searcher and the unvoiced sound-source codebook searcher; a voice reproducer for adding an output of the inverted discrete cosine transformer with pitch period information from the decoder to reproduce a sound source when the output is voiced; and a combining filter for outputting a reproduced voice signal due to outputs of the sound-source reproducer and the spectrum parameter codebook searcher.

[0005]

[Mode for Carrying Out the Invention]

Fig. 1 is a block diagram showing an embodiment of the present invention. In the figure, the upper half shows a transmission-end configuration while the lower half shows a reception-end configuration. On the transmission end, 1 is an LSP (linear spectrum pair) analyzer, 2 is an analyzing filter (inverted filter), 3 is a pitch extractor, 4 is a DCT (discrete cosine transformer), 5 is a switch, 6 is a voiced sound-source

quantizer, 7 is an unvoiced sound-source quantizer, 8 is a voiced sound-source codebook, 9 is an unvoiced sound-source codebook, 10 is an encoder, 11 is an LSP quantizer, 12 is an LSP codebook and 13 is a multiplexer.

[0006]

On the reception end, 14 is a separator for separating a multiplex reception signal inputted through a transmission line into element information, 15 is a switch, 16 is a voiced sound-source codebook searcher, 17 is an unvoiced sound-source codebook searcher, 18 is a voiced sound-source codebook, 19 is an unvoiced sound-source codebook, 20 is a decoder, 21 is an inverted discrete cosine transformer (ICDT), 22 is a sound-source reproducer, 23 is an LSP codebook, 24 is an LSP codebook searcher and 25 is a combining filter.

[0007]

Hereunder, the operation of the embodiment will be explained. First, on the transmission end, the voice signal band-limited and periodically sampled into a digital signal is inputted to the LSP analyzer 1 and analyzed in spectrum parameter. In the LSP analyzer 1, an LSP coefficient is extracted, as a voice spectrum parameter, based on the frame from the voice signal. The spectrum parameter includes, besides LSP coefficients, LPC (line a predictive coding) coefficients, PARCOR (partial auto-correlation) coefficients

and the like. Herein, the case with LSP coefficients is shown as the embodiment.

[0008]

The determined LSP coefficient is sent to the analyzing filter 2 and, at the same time, outputted to the LSP quantizer 11. The analyzing filter 2 calculates an output removed of an LSP coefficient from the voice signal (hereinafter, referred to as voice signal). The LSP coefficient inputted to the LSP quantizer 11 is compared with the LSP coefficient of the LSP codebook holding the LSP coefficient previously determined by calculation according to a training signal, and then quantized. The quantized LSP coefficient is sent to the multiplexer 13.

[0009]

The sound-source signal determined by the analyzing filter 2 is sent to the pitch extractor 3 and DCT (discrete cosine transformer) 4. The pitch extractor 3 detects a pitch period in order to determine a voiced voice and unvoiced voice from the sound-source signal. The determined voiced and unvoiced voice information (voiced/unvoiced information) and the pitch period are sent to the encoder 10 where they are encoded and sent to the multiplexer 13.

[0010]

The DCT 4 transforms the sound-source signal from a time region into a frequency region signal. The sound-source signal transformed into a frequency region signal is sent to

the switch 5 and switch-outputted according to the voiced/unvoiced information determined by the pitch extractor 3. When the sound-source signal is a voiced signal, the sound-source signal is sent to the voiced sound-source quantizer 6 and collated with the content of a voiced sound-source codebook 8 having a frequency component of the voiced sound-source signal previously determined by calculation according to the training signal, being quantized in the corresponding code. When the sound-source signal is an unvoiced signal, the sound-source signal is sent to the unvoiced sound-source quantizer 7 and collated with the content of an unvoiced sound-source codebook 9 having a frequency component of the unvoiced sound-source signal similarly determined according to the training signal, and quantized in the corresponding code. The quantized sound-source signal is sent to the multiplexer 13. The multiplexer 13 multiplexes together the quantized LSP coefficient, the voiced/unvoiced information, the pitch period and the quantized sound-source signal and delivers the same onto the transmission line.

[0011]

On the reception end, the separator 14 receives a multiplex signal from the multiplexer 13 through the transmission line and separates it into a quantized LSP coefficient, encoded voiced/unvoiced information, a pitch period and a quantized sound-source signal. Concerning the

quantized LSP coefficient, a corresponding LSP coefficient is searched out of the LSP codebook 23 same in the content as the transmission-end LSP codebook 12, by the LSP codebook searcher 24. The LSP coefficient searched out by the LSP codebook searcher 24 is sent to the combining filter 25.

[0012]

The voiced/unvoiced information decoded in the decoder 20 is sent to the switch 15 and switched in process route according to any of a voiced and an unvoiced of the quantized sound-source signal. The pitch period decoded in the decoder 20 is sent to the sound-source (residual) reproducer 22.

[0013]

The quantized sound-source signal, when the sound source is a voiced sound, is sent to the voiced sound-source codebook searcher 16 according to the process route switched by the switch 15, and collated with the content of the voiced sound-source codebook 18 having the same content as the transmission-end voiced sound-source codebook 8, being restored into a corresponding voiced sound-source signal.

[0014]

The restored sound-source signal is sent to the IDCT (inverted discrete cosine transformer) 21 and restored into a time region, being sent to the sound-source reproducer 22. The sound-source reproducer 22, when the sound-source signal restored into the time region is a voiced, adds it with the

pitch period information of from the decoder 20 and sends the same to the combining filter 25. The combining filter 25 processes the input sound-source signal by using the LSP coefficient given from the LSP codebook searcher 24 to reproduce and output a voice signal.

[0015]

Although the embodiment explained the case with the scheme using an LSP coefficient as an analysis combining encoding scheme, application is possible to the case with an LPC coefficient as another analysis combining encoding scheme or a scheme using a PARCOR coefficient or the like. voice quality is improved as compared with the conventional simple modeling of sound source.

[0016]

[Effect of the Invention]

As explained in detail above, by carrying out the present invention, improved is the reception reproduced voice quality in voice encoding communication at extremely low rate. Thus the practical effect is great.

[Brief Description of the Drawings]

[Fig. 1] A block diagram showing an embodiment of the present invention.

[Explanation of Reference numerals and Signs]

1 LSP analyzer

2 Analyzing filter

3 Pitch extractor

4 DCT (discrete cosine transformer)

5 Switch (1)

6 Voiced sound-source quantizer

7 Unvoiced sound-source quantizer

8 Voiced sound-source codebook

9 Unvoiced sound-source codebook

10 Encoder

11 LSP quantizer

12 LSP codebook

13 Multiplexer

14 Separator

15 Switch (2)

16 Voiced sound-source codebook searcher

17 Unvoiced sound-source codebook searcher

18 Voiced sound-source codebook

19 Unvoiced sound-source codebook

20 Decoder

21 IDCT (inverted discrete cosine transformer)

22 Sound-source reproducer

23 LSP codebook

24 LSP codebook searcher

25 Combining filter

[Fig. 1]

1 LSP analyzer

2 Analyzing filter

3 Pitch extractor

4 DCT (discrete cosine transformer)

5 Switch (1)

6 Voiced sound-source quantizer

7 Unvoiced sound-source quantizer

8 Voiced sound-source codebook

9 Unvoiced sound-source codebook

10 Encoder

11 LSP quantizer

12 LSP codebook

13 Multiplexer

14 Separator

15 Switch (2)

16 Voiced sound-source codebook searcher

17 Unvoiced sound-source codebook searcher

18 Voiced sound-source codebook

19 Unvoiced sound-source codebook

20 Decoder

21 IDCT (inverted discrete cosine transformer)

22 Sound-source reproducer

23 LSP codebook

24 LSP codebook searcher

25 Combining filter

Input

Output

Transmission line